Our Docket No: 42P17823

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Yoram Gat	Examiner: HUNG, Yubin
Application No: 10/815,896	Art Unit: 2624
Filed: March 31, 2004	Confirmation: 9302
For: Image Segmentation Using Branch and Bound Analysis	

DECLARATION OF INVENTOR YORAM GAT

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir or Madam:

The Applicant respectfully submits a Declaration of Inventor Yoram Gat in connection with the Response filed April 10, 2008.

Declaration of Inventor Yoram Gat

I, Yoram Gat, hereby affirm and declare that:

1. I am one of the two inventors of the invention provided described in the claims of the patent application entitled "Image Segmentation Using Branch and Bound Analysis", filed March 31, 2004, U.S. Patent Application No: 10/815,896 (hereinafter referred to as the "Application"). Intel Corporation, of Santa Clara, California, is the assignee of the Application

2. Horst Haussecker and I are the co-inventors of the above-described patent application and the co-inventors of the subject matter described and claimed therein. The claims of the Application include contributions of mine and of co-inventor Horst Haussecker.

3. I am the author of the article entitled "A Branch-and-Bound Technique for Nano-structure Image Segmentation", provided in the Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop (the "Article"). The Article is attached hereto as Exhibit A.

4. The Article describes certain portions of the invention described in the claims of the Application that were conceived and developed by me. Other elements of such claims include portions of the invention conceived and developed by co-inventor Horst Haussecker.

Exhibit A -- Exhibit A is "A Branch-and-Bound Technique for Nano-structure Image Segmentation", from the Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the above-identified application or an patent issued therefrom.

Respectfully submitted,

Date 4/23, 2008 ______________________
Yoram Gat

EXHIBIT A

A branch-and-bound technique for nano-structure image segmentation

Yoram Gat

Intel research
2200 Mission College Blvd., Santa Clara, CA 95054
yoram.gat@intel.com

Abstract

Images of nano-structures are often noisy. On the other hand, in many settings there is quite a lot of model knowledge regarding the observed structures. This paper proposes a method for segmenting an image using a geometric model of the the observed structure. The resulting segmentation is guaranteed to be globally optimal, for an explicitly specified score function. This property provides a great deal of robustness to the algorithm.

The algorithm presented explores a pre-defined space of segmentations using a branch-and-bound algorithm. It eliminates those parts of the space that are provably poor and explores in further detail the more promising parts of the space.

An example of a segmentation that can be obtained in this way is a straight line segmentation of an image into 2 regions that minimizes the intensity variation within the regions. Results showing extraction of specific nano-structures are presented.

A trivial variation on the algorithm can find a maximum a-posteriori probability estimate of the segmentation when there exists an a-priori distribution over the segmentations and the objective function is interpreted as the likelihood of the image given the segmentation.

1 Introduction

Image segmentation is a widely explored problem in image processing. It is the operation of partitioning an image into regions which correspond to meaningful objects. In the context of nano-vision, this often involves finding relatively simple, well-defined shapes such as lines, circles or rectangles in noisy images.

One of the common methods for segmentation is region-based segmentation. Region-based segmentation methods attempt to find in the image regions which are homogeneous according to some criterion, such as color, intensity or texture. The objective is to find regions that are both homogeneous and possess good shape properties - such as connectedness and edge smoothness.

The method presented here handles both parts of the objective explicitly. The shape properties of the segmentations under consideration are defined, and among all segmentations with those properties, the one which maximizes an explicit homogeneity criterion is obtained.

The guaranteed optimal segmentation is obtained using a branch-and-bound (B&B) search approach. Such an approach has been used before in image analysis for object recognition ([2], [3]). Discussion of the relationship of this approach and other methods related to the one proposed is presented in section 2.

Informally, the setup is as follows: A set of possible segmentations is being considered. Say, the set of all straight-line segmentations of the image into two parts. Given a particular segmentation, an objective function gives the cost of the segmentation, with lower costs preferable.

It is assumed that a lower bound of the cost over sets of segmentations can be calculated. That is, given a set of segmentations, a value can be determined which is less than or equal to the cost of the best (least costly) segmentation in the set. While in general the lower bound will not be equal to the cost of the best segmentation in the set, it is assumed that for small sets, the bound is tight.

This setup is formalized in Section 3.

The B&B algorithm that determines the globally best segmentation works as follows: It maintains a priority queue, where the elements in the queue are sets of segmentations and their priority is equal to the lower bound of the cost of the best segmentation in the set (the smaller the bound, the higher the priority).

The set with highest priority in the queue is removed, and split into parts. Each of those parts is then inserted to the queue. The algorithm stops when the set with the highest priority contains only one segmentation (or some other stopping criterion is met).

A more formal description of the algorithm is presented in Section 4.

Section 5 describes interpretations of the setup in the statistical terms and Section 6 presents a few examples of the way the algorithm performs with natural and synthetic images.

Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'03)
1063-6919/03 $ 17.00 © 2003 IEEE

COMPUTER SOCIETY

Docket No.: 42P17823
Application No.: 10/815,896

2 Relation to previous work

The generalized Hough transform (HT) (e.g. [1]) is a generic method for locating parameterized families of objects in an image. It is very similar in approach to the method proposed here in the sense that both have an explicit representation of the shape that is to be located. The objective function of the Hough transform is based on votes accumulated from pixels, where each pixel is seen as voting for some parameter values. In the setup here the objective function is unrestricted in its form.

A major difference between the method proposed here and HT is that HT is implemented by quantizing the parameter space and evaluating the objective function for each point in the quantized parameter space, while the B&B method described here works in continuous space and uses bounds to handle regions of the parameter space without evaluating dense sets of points in those regions.

The Hough transform therefore has to strike a balance between quantizing too coarsely and incurring approximation errors and quantizing too finely and incurring high computational cost. The B&B method avoids approximation errors altogether and has the potential to arrive quickly at an exact solution when the objective function is well behaved.

A different approach to image analysis and segmentation that is related to the one presented here is that of Mount, Netanyahu and Le Moigne ([3]). Following [2] they consider the problem of matching point sets using partial Hausdorff distance. They suggest a B&B algorithm to obtain good matches. The approach here is very similar, but the objective function considered here is not restricted to being the partial Hausdorff distance, and the focus is not on point matching.

Lastly, another work that can be contrasted with the approach here is that of Shi and Malik ([5]). As is done here, they address the segmentation problem as an optimization problem. Rather than having a pre-defined set of possible edge shapes, Shi and Malik have the shape criterion as part of the objective function. This allows for much greater variety of shapes of segment edges, but reduces the control the user has over desired edge shapes. For example, it would be difficult to enforce straight or circular edges, if these are desirable due to some a-priori considerations (as is often the case in the context of nano-vision).

3 The setup

The image is implicitly assumed fixed throughout the discussion. The image data may include any pre-processing of the raw image such as texture features.

Let $\mathcal{P}$ be the set of pixels in the image:

$$\mathcal{P} = \{(i,j) : i = 1,\ldots,n_x,\ j = 1,\ldots,n_y\}.$$

Example For illustration, the general description of the setup and the algorithm is accompanied by a concrete example. The example is as follows: Each pixel $p \in \mathcal{P}$ is associated with a scalar $v(p)$ - its intensity value, say. Straight line, 2-part segmentations are considered. Among all such segmentations, the one which minimizes the sum of absolute value deviations of the values v from the segment median is sought.

The details of the algorithm for this particular segmentation problem are presented alongside the generic description of the algorithm.

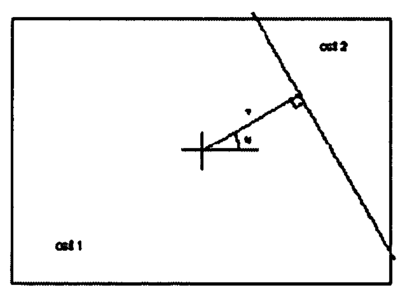
Let $\mathcal{S}$ be a parameterized family of K-part segmentations. For example, the set of all straight lines is a family of 2-part segmentations. Formally, any $s \in \mathcal{S}$ is a map from $\mathcal{P}$ into the set of cell indices $\{1,\ldots,K\}$.

Example The space of straight line segmentations can be parameterized as follows:

$$\mathcal{S}_{\text{LIN}} = \{s_{\theta,r} : \theta \in [0,2\pi], r \in [0, r_{max}]\},$$

where $r_{max} = \frac{1}{2}\sqrt{n_x^2 + n_y^2}$ and

$$s_{\theta,r}(i,j) = \begin{cases} 1 & \text{if } (i - \frac{n_x}{2})\cos\theta + \\ & \quad (j - \frac{n_y}{2})\sin\theta < r \\ 2 & \text{otherwise} \end{cases}.$$



Let l be the segmentation loss function, scoring each segmentation according to some (fixed) criterion:

$$l : \mathcal{S} \to \mathbf{R}.$$

Example The example segmentation loss function is

$$l(s) = l_1(s) + l_2(s),$$

Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW03)
1063-6919/03 $ 17.00 © 2003 IEEE

Docket No.: 42P17823
Application No.: 10/815,896

where

$$l_i(s) = \sum_{p: s(p)=i} |v(p) - \tilde{v}_{s,i}|,$$

with $\tilde{v}_{s,i}$ being the median value of the set $\{v(p) : s(p) = i\}$.

The algorithm described here is aimed at finding a segmentation that minimizes the objective function l over the set of segmentations $\mathcal{S}$. It assumes the existence of the following:

- A refining function R -

$$R : 2^{\mathcal{S}} \to 2^{2^{\mathcal{S}}}.$$

This function breaks up sets of segmentations into smaller sets by mapping a set of segmentations $S \subseteq \mathcal{S}$ into a set of segmentation sets $S_1, \ldots, S_l$ which form a partition of S. The function needs only be defined for those sets which can be obtained by refining $\mathcal{S}$.

Example For notational ease define

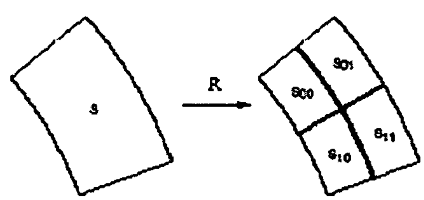
$$S_{[\theta_0,\theta_1]\times[r_0,r_1]} = \{s_{\theta,r} : \theta \in [\theta_0,\theta_1], r \in [r_0,r_1]\},$$

giving $\mathcal{S}_{\text{LIN}} = S_{[0,2\pi]\times[0,r_{max}]}$.

The refining function for the example setup is defined on, and maps to sets of, set of segmentations of the form $S_{[\theta_0,\theta_1]\times[r_0,r_1]}$.
It is defined as follows:

$$R(S_{[\theta_0,\theta_1]\times[r_0,r_1]}) = \Big\{S_{[\theta_0,\theta']\times[r_0,r']},$$
$$S_{[\theta_0,\theta']\times[r',r_1]}, S_{[\theta',\theta_1]\times[r_0,r']}, S_{[\theta',\theta_1]\times[r',r_1]}\Big\},$$

where $\theta' = (\theta_0+\theta_1)/2$ and $r' = (r_0+r_1)/2$.



- A partial segmentation function M -

$$M : 2^{\mathcal{S}} \times \mathcal{P} \to 2^{\{1,\ldots,K\}}.$$

The function M must have the following property: For any $p \in \mathcal{P}$,

$$\{s(p) : s \in S\} \subseteq M(S,p).$$

That is, the function M provides a superset of all the cells that a pixel may be mapped to by a set of segmentations. As is the case with the refining function R, the function M needs only to be defined when the first argument is a subset of $\mathcal{S}$ that can be obtained by refining $\mathcal{S}$.

Example A partial segmentation function for linear segmentations is defined as follows: Fix $\theta_0, \theta_1, r_0, r_1$ and let

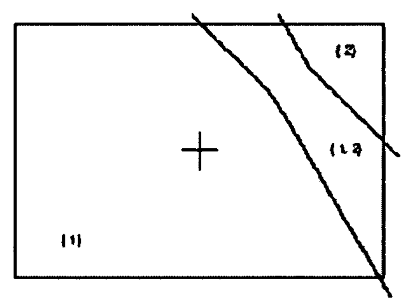
$$t_0 = \min\Big\{(i - \frac{n_x}{2})\cos\theta + (j - \frac{n_y}{2})\sin\theta : \theta \in [\theta_0,\theta_1]\Big\},$$

and

$$t_1 = \max\Big\{(i - \frac{n_x}{2})\cos\theta + (j - \frac{n_y}{2})\sin\theta : \theta \in [\theta_0,\theta_1]\Big\}.$$

Let

$$M(S_{[\theta_0,\theta_1]\times[r_0,r_1]}, (i,j)) = \begin{cases} \{1\} & \text{if } t_1 < r_0, \\ \{2\} & \text{if } t_0 > r_1, \\ \{1,2\} & \text{otherwise} \end{cases}.$$



- A partial segmentation loss function B -

$$B : \left(2^{\{1,\ldots,K\}}\right)^{\mathcal{P}} \to \mathbb{R}.$$

This function maps a partial segmentation of the pixels in the image to a loss value. This loss value must be less than or equal to all loss values associated with any segmentation that is consistent with the partial segmentation. That is if T is a map from pixels to subsets of $\{1,\ldots,K\}$, then

$$B(T) \leq \inf\{l(s) : s \in \mathcal{S}, \text{for all } p \in \mathcal{P}, s(p) \in T(p)\}.$$

Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'03)
1063-6919/03 $ 17.00 © 2003 IEEE

Docket No.: 42P17823
Application No.: 10/815,896

Example A partial segmentation loss function for the loss function of the example is

$$B(T) = B_1(T) + B_2(T),$$

where

$$B_i(T) = \sum_{p:T(p)=(i)} |v(p) - \bar{v}_{T,i}|,$$

with $\bar{v}_{T,i}$ being the median value of the set

$$\left\{ v(p) : T(p) = (\ i\ ) \right\}.$$

4 A branch-and-bound algorithm

The branch-and-bound algorithm which is used to search for an optimal segmentation is a standard A* procedure ([4]). The procedure uses a priority queue to store 'states'. A state may be a terminal state, which means that when it reaches the head of the queue the search has terminated. Otherwise the state is non-terminal, and so when it reaches the head of the queue, it is removed from the queue and all the states that are immediately reachable from that state are inserted to the queue. Each state in the queue has a priority score associated with it, and the state with the lowest priority score (i.e., highest priority) is considered to be at the head of the queue.

In the present setup the states are subsets of the segmentation space S. For any non-terminal state, S, the states which are immediately reachable from S are given by $R(S)$.

Using M and B, the priority function L,

$$L : 2^S \rightarrow \mathbf{R},$$

is defined as follows:

$$L(S) = B(T),$$

where the partial segmentation T is given by $T(p) = M(S,p)$. By construction, the function L provides a lower bound for the loss of any segmentation which is an element of the set which is its argument:

$$L(S) \leq \inf\{l(s) : s \in S\}.$$

The tightness of $L(S)$ as a bound to $\inf\{l(s) : s \in S\}$ is a major factor in determining the efficiency of the segmentation algorithm. For the algorithm to converge to the optimal segmentation it is necessary that the following limit holds for any sequence of sets $S_0, S_1, \ldots$, generated by refining S:

$$\lim_{i \to \infty} L(S_i) = l(\cap_{i=0}^{\infty} S_i).$$

A state is terminal if all the segmentations in the segmentation set agree on all pixels, or if some other termination condition is met. The search is initialized by inserting to the queue the state which corresponds to the whole segmentation space S. The algorithm is therefore:

- **Initialization** Start with empty priority queue. Insert into the queue the state S, with the corresponding priority score $L(S)$.
- **Extract** Extract the state at the head of the queue - S. Remove it from the queue. If it is a terminal state - stop, output S as the solution to the optimization problem.
- **Insert** Insert to the queue the members of the set $R(S) = \{S_1, \ldots, S_k\}$, with the priority scores $L(S_1), \ldots, L(S_k)$ correspondingly.
- **Iterate** Go back to step Extract.

5 Frequentist and Bayesian statistical interpretation

If the loss function $l(s)$ can be interpreted as the negative of the log-likelihood of the image given the segmentation s, then the search for minimum loss can be interpreted as maximum likelihood estimation.

Trivially, the same algorithm can be used to obtain a Bayesian maximum a-posteriori probability (MAP) estimate of the segmentation. The loss associated with a segmentation would then be interpreted as the sum of the negative of the log-probability of the image given the segmentation and the negative of the log of the a-priori probability of the segmentation.

In addition to obtaining the MAP estimate, a credible region can be obtained, by employing a similar algorithm to the one presented. This is a region of the segmentation space that a supports a proscribed part of the probability in the a-posteriori distribution of the segmentation.

In order to do this, upper bounds on the loss function, as well as lower bounds, need to be evaluated.

6 Results for test images

This section shows the output of the segmentation algorithm when used to extract features from a focused ion beam imaging (FIB) tool image. The features are copper wires on a background of silicon. The segmentation space considered for each image was constructed specifically to extract the expected structure: an L shaped corner, a T-junction or a wire termination.

Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW03)
1063-6919/03 $ 17.00 © 2003 IEEE

Docket No.: 42P17823
Application No.: 10/815,896

Edge shape	# of cells	# of parameters	# of $L(S)$ evaluations	Runtime (secs.)
L-corner	2	4	1,408	12.4
T-junction	2	4	4,000	6.9
Wire tip	2	3	1,184	4.4

Table 1: Experiment details. All experiments were run on a 2.4GHz Pentium 4 desktop system, with 2GB RAM. Image sizes vary between 172 x 115 and 352 x 313 pixels.

The objective function for all the test cases was:

$$\sum_{i\in\{1,2\}} \sum_{p:t(p)=i} |v(p) - \bar{v}_i| ,$$

where i enumerates the different cells, and $\bar{v}_i$ is the median intensity value of the pixels in cell i.

The partial segmentation loss function B used was as follows:

$$B(T) = \min_{a_1,a_2} \left(B^1(T, a_1) + B^2(T, a_2) + B^{1,2}(T, a_1, a_2)\right) ,$$

where

$$B^1(T, a_1) = \sum_{p:T(p)=\{1\}} |v(p) - a_1| ,$$

$$B^2(T, a_2) = \sum_{p:T(p)=\{2\}} |v(p) - a_2| ,$$

and

$$B^{1,2}(T, a_1, a_2) = \sum_{p:T(p)=\{1,2\}} \min(|v(p) - a_1| , |v(p) - a_2|).$$

This function can be evaluated in time that is proportional to the number of possible values of color intensity, i.e., 256.

The results are in Figure 1.

Table 1 shows some details regarding the experiments.

To demonstrate the robustness of the algorithm to noise, it was also run on a synthetic image that contained an edge obscured by noise. The algorithm is able to recover the edge (Figure 2) even when the noise is high enough to challenge the human eye.

For the synthetic color image, the objective function used was the sum of the three objective values obtain by evaluating the single channel objective function for each of the three color channels.

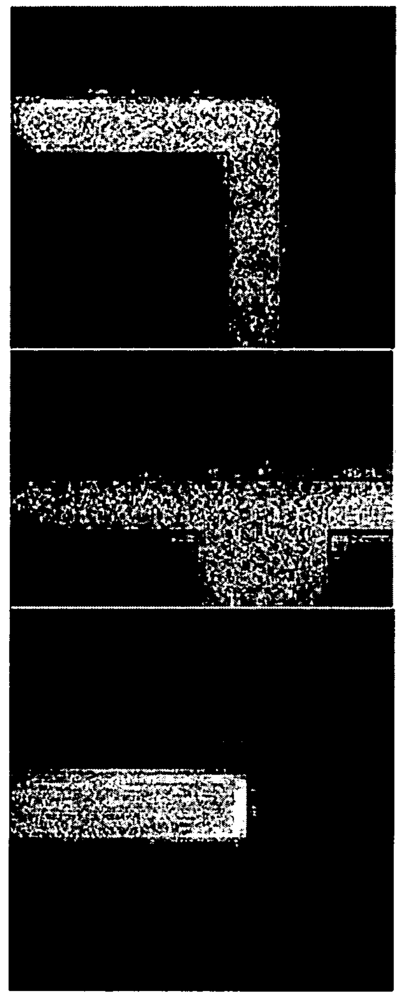


Figure 1: Various nano-structured segmented out of FIB images

7 Conclusions and future work

The work presented here demonstrated that segmentation of nano-structure images can be obtained by optimizing a simple objective function over a space of partitions with pre-defined shape properties. With computing power available

Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW03)
1063-6919/03 $ 17.00 © 2003 IEEE

Docket No.: 42P17823
Application No.: 10/815,896

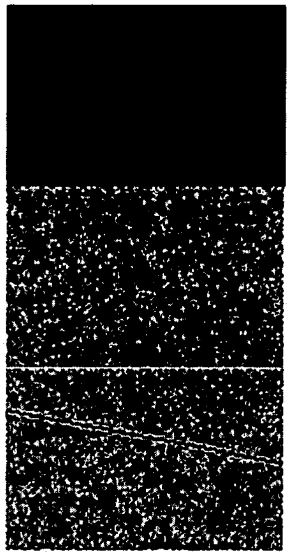


Figure 2: (Top to bottom) Ground truth edge, ground truth + noise = input, output

today, finding solutions that are globally optimal is feasible but cannot be done in real time.

A useful extension of the algorithm that was presented would be using a similar approach to obtain multiple local optima, in addition to the global optimum. This would enable picking out from an image multiple objects with similar shape characteristics.

Another interested line of research would be using physics based objective functions instead of the heuristic functions used in the present paper.

Acknowledgments

I thank Alan Pobanz for helping acquire the FIB images and Jean-Yves Bouguet and Ara Nefian for reading earlier versions of this paper and providing valuable comments.

References

[1] Davies, E.R. (1997). *Machine vision*, Academic Press.

[2] Huttenlocher, D.P., Klanderman, G.A. and Rucklidge W.J. (1993). Comparing Images Using the Hausdorff Distance, *IEEE transactions on pattern analysis and machine intelligence* **15** (9).

[3] Mount D.M., Netanyahu N.S. and Le Moigne J. (1998). Efficient Algorithms for Robust Feature Matching. *Pattern recognition* **32** (13).

[4] Russel, S. J. and Norvig, P. (1995). *Artificial intelligence*, Prentice Hall.

[5] Shi, J. and Malik, J. (2000). Normalized cuts and image segmentation, *IEEE transactions on pattern analysis and machine intelligence* **22** (8).

Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'03)
1063-6919/03 $ 17.00 © 2003 IEEE

Docket No.: 42P17823
Application No.: 10/815,896